

Confusion matrix

| | | Predicted Class | | |
|--------------|----------|--------------------------------------------|------------------------------------------------------------|----------------------------------------------------------|
| | | Positive | Negative | |
| Actual Class | Positive | True Positive (TP) | False Negative (FN) Type II Error | Sensitivity (RECALL) $\frac{TP}{(TP + FN)}$ |
| | Negative | False Positive (FP) Type I Error | True Negative (TN) | Specificity $\frac{TN}{(TN + FP)}$ |
| | | Precision $\frac{TP}{(TP + FP)}$ | Negative Predictive Value $\frac{TN}{(TN + FN)}$ | Accuracy $\frac{TP + TN}{(TP + TN + FP + FN)}$ |



Confusion matrix

| | Actual Class: Positive | Actual Class: Negative |
|--------------------|--------------------------------------|--------------------------------------|
| Predicted Positive | TP (True Positive) | FP (False Positive) |
| Predicted Negative | FN (False Negative) | TN (True Negative) |

Sensitivity (Recall) = $TP / (TP + FN)$

Specificity = $TN / (TN + FP)$

Precision = $TP / (TP + FP)$

Negative Predictive Value = $TN / (TN + FN)$

Accuracy = $(TP + TN) / (TP + TN + FP + FN)$

F1 Score = $2 * (Precision * Recall) / (Precision + Recall)$



Accuracy / Dokładność

- Accuracy = ogólna trafność klasyfikatora
- Stosunek ilości poprawnie zaklasyfikowanych obserwacji do ilości wszystkich obserwacji

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN}$$



Precision - spośród wszystkich przewidzianych "tak", ile faktycznie było "tak"?

- Precision = dokładność klasyfikacji w obrębie prognozowanej klasy „positive” (jakość przewidzianych "tak")
- Inne nazwy: Positive Predictive Value (PPV), Precyzja

$$\text{Precision} = \frac{TP}{TP + FP}$$



Sensitivity (Recall) - spośród wszystkich rzeczywistych „tak”, ile model poprawnie wykrył jako „tak”?

- Sensitivity = dokładność klasyfikacji w obrębie rzeczywistej klasy „positive” (kompletność wykrycia rzeczywistych "tak")
- Inne nazwy: True Positive Rate (TPR), Recall, Czułość / wrażliwość

$$\text{Recall} = \frac{TP}{TP + FN}$$



Specificity - spośród wszystkich rzeczywistych "nie", ile model poprawnie wykrył jako "nie"?

- Specificity = dokładność klasyfikacji w obrębie rzeczywistej klasy „negative”
(kompletność wykrycia rzeczywistych "nie")
- Inne nazwy: True Negative Rate (TNR),
Specyficzność

$$\text{Specificity} = \frac{TN}{TN + FP}$$



Negative Predictive Value - spośród wszystkich przewidzianych „nie”, ile faktycznie było „nie”?

- NPV = dokładność klasyfikacji w obrębie prognozowanej klasy „negative” (jakość przewidzianych "nie")
- Inne nazwy: Negative Predictive Value (NPV), Precyzja negatywna

$$NPV = \frac{TN}{TN + FN}$$



Porównanie metryk

| Metryka | Pytanie, na które odpowiada |
|----------------------------------------|-----------------------------------------------------------------------------------------------------------------------|
| Precision (PPV) | Jeśli model przewiduje „tak”, czy przypadek rzeczywiście jest „tak”? |
| Recall (Sensitivity, TPR) | Jeśli przypadek rzeczywiście jest „tak”, czy model go poprawnie rozpozna jako „tak”? |
| Specificity (TNR) | Jeśli przypadek rzeczywiście jest „nie”, czy model go poprawnie rozpozna jako „nie”? |
| Negative Predictive Value (NPV) | Jeśli model przewiduje „nie”, czy przypadek rzeczywiście jest „nie”? |
| F1 Score | Jak dobrze model równoważy poprawność przewidywań „tak” (Precision) i ich kompletność (Recall)? |
| Accuracy | Jaki procent wszystkich przypadków został poprawnie sklasyfikowany (zarówno „tak”, jak i „nie”)? |
| AUC (Area Under ROC Curve) | Jak dobrze model rozróżnia wszystkie przypadki „tak” od wszystkich przypadków „nie” przy różnych progach decyzyjnych? |



F-measure

- F-measure = wykorzystuje średnią harmoniczną precyzji (precision) i czułości (sensitivity, recall)
- Inne nazwy: Miara F1, F1-measure, F1-score

Miara F (F-measure)

$$F_{\beta} = (1 + \beta^2) \cdot \frac{\text{precision} \cdot \text{recall}}{\beta^2 \cdot \text{precision} + \text{recall}}$$

Miara F1 (F1-measure, F1-score) – w powyższym wzorze $\beta=1$

$$F = 2 \cdot \frac{\text{precision} \cdot \text{recall}}{\text{precision} + \text{recall}}$$



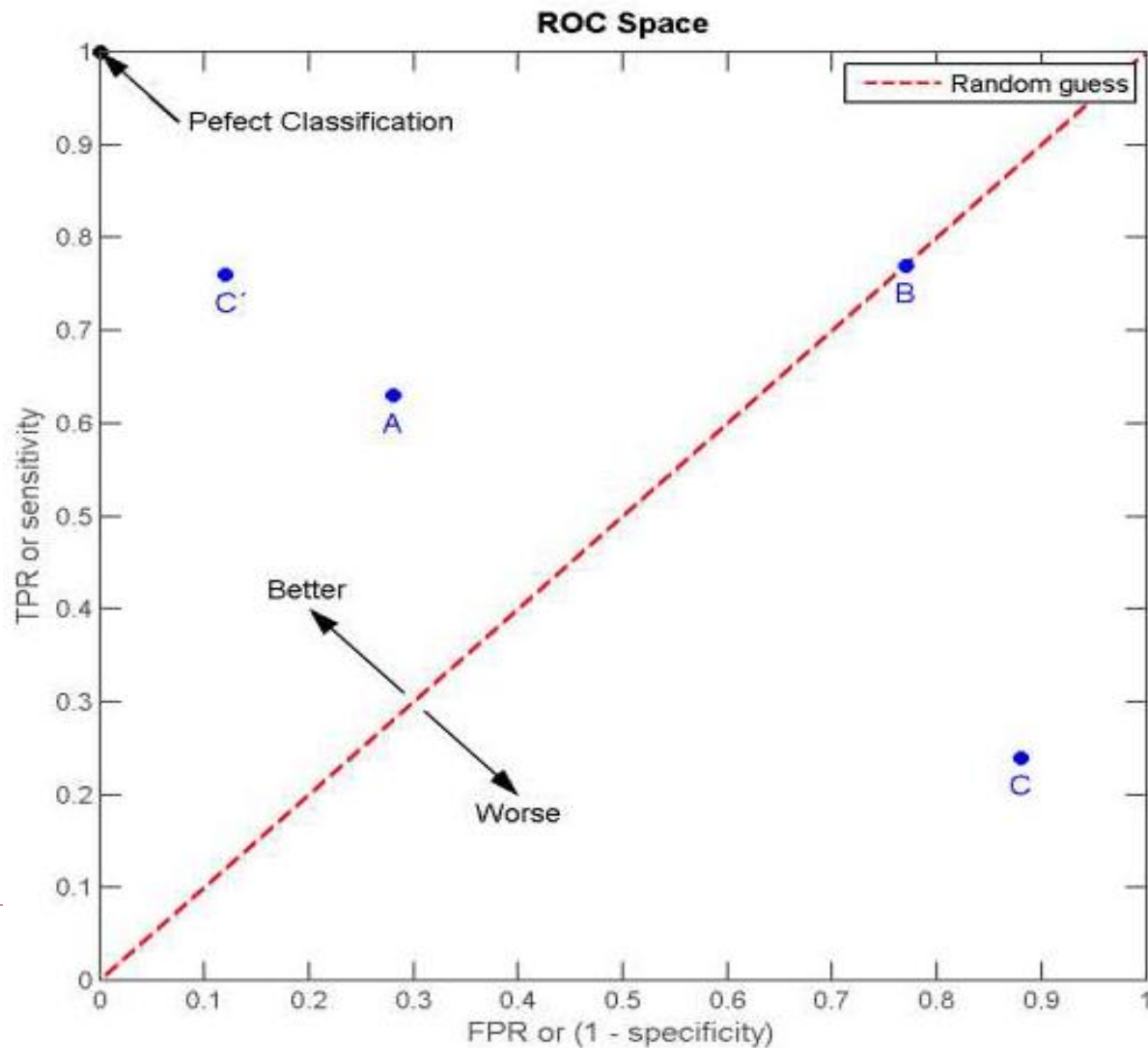
Krzywa ROC

- Receiver Operating Characteristics
- wykres charakterystyczny dla danego klasyfikatora, bada stosunek wielkości TP do FP
- narzędzie pomagające w określeniu wydajności klasyfikatorów binarnych
- zawiera wiele punktów dla różnych wartości progu decyzyjnego klasyfikatora
- Fall-out (False Positive Rate) = 1-Specificity =:

$$\text{FPR} = \frac{\text{False Positives (FP)}}{\text{False Positives (FP)} + \text{True Negatives (TN)}}$$



Krzywa ROC



Do przeczytania o miarach

- Accuracy, precision, recall, F1 – co to za czary?

<https://jakbadacdane.pl/accuracy-precision-recall-f1-co-to-za-czary/>

- Precision, recall i F1 – miary oceny klasyfikatora

<https://ksopyla.com/machine-learning/precision-recall-f1-miary-oceny-klasyfikatora/>

- What is Confusion Matrix and Advanced Classification Metrics?

<https://manisha-sirsat.blogspot.com/2019/04/confusion-matrix.html>

- Performance Metrics for Classification problems in ML

<https://medium.com/@MohammedS/performance-metrics-for-classification-problems-in-machine-learning-part-i-b085d432082b>

